

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number: 09/755,320
Source: IFW16
Date Processed by STIC: 01/26/2006

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 01/26/2006

PATENT APPLICATION: US/09/755,320

TIME: 10:10:26

Input Set : N:\Crf3\RULE60\09755320.raw.txt

Output Set: N:\CRF4\01262006\I755320.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Gorski, David H.

6 Walsh, Kenneth

8 (ii) TITLE OF INVENTION: Growth Arrest Homeobox Gene

10 (iii) NUMBER OF SEQUENCES: 19

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: Calfee, Halter, and Griswold

14 (B) STREET: 800 Superior Avenue

15 (C) CITY: Cleveland

16 (D) STATE: Ohio

17 (E) COUNTRY: U.S.A.

18 (F) ZIP: 44114-2688

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Floppy disk

22 (B) COMPUTER: IBM PC compatible

23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

24 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

26 (vi) CURRENT APPLICATION DATA:

C--> 27 (A) APPLICATION NUMBER: US/09/755,320

C--> 28 (B) FILING DATE: 05-Jan-2001

W--> 34 (C) CLASSIFICATION: 435

31 (vii) PRIOR APPLICATION DATA:

32 (A) APPLICATION NUMBER: US/08/203,532

33 (B) FILING DATE: 24-Feb-1994

36 (viii) ATTORNEY/AGENT INFORMATION:

37 (A) NAME: Golrick, Mary E.

38 (B) REGISTRATION NUMBER: 34829

39 (C) REFERENCE/DOCKET NUMBER: 22311/00114

41 (ix) TELECOMMUNICATION INFORMATION:

42 (A) TELEPHONE: (216) 622-8200

43 (B) TELEFAX: (216) 241-0816

44 (C) TELEX: 980499

47 (2) INFORMATION FOR SEQ ID NO: 1:

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 2244 base pairs

51 (B) TYPE: nucleic acid

52 (C) STRANDEDNESS: both

53 (D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: cDNA

57 (iii) HYPOTHETICAL: NO

59 (iv) ANTI-SENSE: NO

62 (ix) FEATURE:

RAW SEQUENCE LISTING

DATE: 01/26/2006

PATENT APPLICATION: US/09/755,320

TIME: 10:10:26

Input Set : N:\Crf3\RULE60\09755320.raw.txt

Output Set: N:\CRF4\01262006\I755320.raw

```

63      (A) NAME/KEY: CDS
64      (B) LOCATION: 197..1108
65      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
69 GTCAAGTGT TATACGTGCA GGAGACTGGC CGCTCGGCTC AGGACTGGGA TTAGCGGGCT      60
71 CTGCTCAAAC CCGCGCGGCT TTTACATTAG GAGTGAGTGG GGGAGAGTCC TAGGATTTCT      120
73 AGTGAAAAGT GACAGCGCTT GGTGGACTTT GGGACCTTCG TGAAGTCTTC TGCTTGGAAG      180
75 CTGAGACTTG CATGCC ATG GAA CAC CCC CTC TTT GGC TGC CTG CGC AGC      229
76      Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser
77      1          5          10
79 CCC CAC GCC ACA GCG CAA GGC TTG CAC CCC TTC TCG CAG TCT TCT CTG      277
80 Pro His Ala Thr Ala Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu
81      15          20          25
83 GCC CTC CAT GGA AGA TCT GAC CAC ATG TCC TAC CCC GAA CTC TCC ACA      325
84 Ala Leu His Gly Arg Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr
85      30          35          40
87 TCT TCC TCG TCT TGC ATA ATC GCG GGA TAC CCC AAT GAG GAG GGC ATG      373
88 Ser Ser Ser Ser Cys Ile Ile Ala Gly Tyr Pro Asn Glu Glu Gly Met
89      45          50          55
91 TTT GCC AGC CAG CAT CAC AGG GGG CAC CAC CAC CAC CAC CAC CAC      421
92 Phe Ala Ser Gln His His Arg Gly His His His His His His His
93 60      65      70      75
95 CAT CAC CAC CAC CAG CAG CAG CAG CAC CAG GCT CTG CAA AGC AAC TGG      469
96 His His His His Gln Gln Gln Gln His Gln Ala Leu Gln Ser Asn Trp
97      80      85      90
99 CAC CTC CCG CAG ATG TCC TCC CCG CCA AGC GCG GCC CGG CAC AGC CTT      517
100 His Leu Pro Gln Met Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu
101      95      100      105
103 TGC CTG CAG CCT GAT TCC GGA GGG CCC CCG GAG CTG GGG AGC AGC CCT      565
104 Cys Leu Gln Pro Asp Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro
105      110      115      120
107 CCG GTC CTG TGC TCC AAC TCT TCT AGC CTG GGC TCC AGC ACC CCG ACC      613
108 Pro Val Leu Cys Ser Asn Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr
109      125      130      135
111 GGA GCC GCG TGC GCA CCA AGG GAT TAT GGC CGT CAA GCG CTG TCA CCC      661
112 Gly Ala Ala Cys Ala Pro Arg Asp Tyr Gly Arg Gln Ala Leu Ser Pro
113 140      145      150      155
115 GCA GAA GTG GAG AAG AGA AGT GGC AGC AAA AGA AAA AGC GAC AGT TCA      709
116 Ala Glu Val Glu Lys Arg Ser Gly Ser Lys Arg Lys Ser Asp Ser Ser
117      160      165      170
119 GAT TCC CAG GAA GGA AAT TAC AAG TCA GAA GTG AAC AGC AAA CCT AGG      757
120 Asp Ser Gln Glu Gly Asn Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg
121      175      180      185
123 AGG GAA AGA ACA GCT TTC ACC AAA GAG CAA ATC AGA GAA CTT GAG GCA      805
124 Arg Glu Arg Thr Ala Phe Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala
125      190      195      200
127 GAG TTC GCC CAT CAT AAC TAT CTG ACC AGA CTG AGA AGA TAT GAG ATA      853
128 Glu Phe Ala His His Asn Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile
129      205      210      215
131 GCG GTG AAC CTA GAC CTC ACT GAA AGA CAG GTG AAA GTG TGG TTC CAG      901

```

RAW SEQUENCE LISTING

DATE: 01/26/2006

PATENT APPLICATION: US/09/755,320

TIME: 10:10:26

Input Set : N:\Crif3\RULE60\09755320.raw.txt

Output Set : N:\CRF4\01262006\I755320.raw

```

132 Ala Val Asn Leu Asp Leu Thr Glu Arg Gln Val Lys Val Trp Phe Gln
133 220                225                230                235
135 AAC AGG AGA ATG AAG TGG AAG CGG GTC AAG GGG GGA CAA CAA GGA GCT      949
136 Asn Arg Arg Met Lys Trp Lys Arg Val Lys Gly Gly Gln Gln Gly Ala
137                240                245                250
139 GCA GCC CGA GAA AAG GAA CTG GTG AAT GTG AAA AAG GGA ACA CTT CTT      997
140 Ala Ala Arg Glu Lys Glu Leu Val Asn Val Lys Lys Gly Thr Leu Leu
141                255                260                265
143 CCA TCA GAG CTG TCA GGA ATT GGT GCA GCC ACC CTC CAG CAG ACA GGG      1045
144 Pro Ser Glu Leu Ser Gly Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly
145                270                275                280
147 GAC TCA CTA GCA AAT GAC GAC AGT CGC GAT AGT GAC CAC AGC TCT GAG      1093
148 Asp Ser Leu Ala Asn Asp Asp Ser Arg Asp Ser Asp His Ser Ser Glu
149                285                290                295
151 CAC GCA CAC TTA TGATACATAC AGAGACCAGC TCCGTTCTCA GGAAAGCACC      1145
152 His Ala His Leu
153 300
155 ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG      1205
157 CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC      1265
159 ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA      1325
161 AATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTTGG CTTGCACTGA      1385
163 AAATTAAATT GCTACCAAGA GCAAACATCG TAAGACATTT TGAATCAAGT TGTCTCCAGA      1445
165 GTGAAGATGT TATAGAAATG CTTTGAACAT TCCAGTTGTA CCAGGTCATG TGTGTGACAC      1505
167 TGGGCAGGTA TTTGCTTTTG CTTGCACTGA AACTTAAACT GCTATCAAGT TAACCCATGA      1565
169 AATAGTTTAT CTTGAACAGC CACAGTGCCT GAAATCACCA AGTGGATATA AAATGAACTG      1625
171 AAATTCTGTA TATATTACTC CTAAGTCATT TTCCTGTCTT CACTAATTTT AGCAAATGCA      1685
173 TTCATATTAG CTGATGAAAA TAGGCTTTCC CGTGGACAAA TGCAGCCAGC TTCTTGATT      1745
175 TTTATACATT TTTTGTGTCAG TCAGAGACAT CAGTATGTGC TTAAGTGTGT TCAAGTAGAG      1805
177 GAAATGCAGT AGAGTCTGAT AGGACATATT CTTGGTACCA CAGACAAAAC AAATCTTCTG      1865
179 TTGCATTGAC TATCAACTGC TGCAGATACA TTAGAGAACA CACCTAGCCC CCCTCCAGCC      1925
181 TCCCTCTGTT ATCGCTCGAA GACATTAGCG TCATAGGCAA GTAGTTACCT TGCCAAATGA      1985
183 GTCTTGTGTG GCAGATGTCT GATTTTGTAT CTTTAAACTG TTAATGGTAT GTGTCTGCTT      2045
185 CAGTTAACAG GGAAAAAGAT TTCTTCCTCA TTGTTTATGA TACAAAACCC AAGTGCCAAA      2105
187 CAAAGCTAGT TCTTCAAGGG ATAGATGAGA AACTGAATGT CTGACAAGTA GACTCAGCGA      2165
189 AAATACATTA TTTTTCAGAG GCTGTGTATT CATGCAGTAC AAGTCCTTGT ATTTTGTAAA      2225
191 AAAAAAAGTT AAATAAATG      2244
194 (2) INFORMATION FOR SEQ ID NO: 2:
196     (i) SEQUENCE CHARACTERISTICS:
197         (A) LENGTH: 303 amino acids
198         (B) TYPE: amino acid
199         (D) TOPOLOGY: linear
201     (ii) MOLECULE TYPE: protein
203     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
205 Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser Pro His Ala Thr Ala
206  1          5          10          15
208 Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu Ala Leu His Gly Arg
209          20          25          30
211 Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr Ser Ser Ser Ser Cys
212          35          40          45

```

RAW SEQUENCE LISTING

DATE: 01/26/2006

PATENT APPLICATION: US/09/755,320

TIME: 10:10:26

Input Set : N:\Crif3\RULE60\09755320.raw.txt

Output Set: N:\CRF4\01262006\I755320.raw

```

214 Ile Ile Ala Gly Tyr Pro Asn Glu Glu Gly Met Phe Ala Ser Gln His
215      50                      55                      60
217 His Arg Gly His His His His His His His His His His His His Gln
218 65                      70                      75                      80
220 Gln Gln Gln His Gln Ala Leu Gln Ser Asn Trp His Leu Pro Gln Met
221      85                      90                      95
223 Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu Cys Leu Gln Pro Asp
224      100                      105                      110
226 Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro Pro Val Leu Cys Ser
227      115                      120                      125
229 Asn Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr Gly Ala Ala Cys Ala
230      130                      135                      140
232 Pro Arg Asp Tyr Gly Arg Gln Ala Leu Ser Pro Ala Glu Val Glu Lys
233 145                      150                      155                      160
235 Arg Ser Gly Ser Lys Arg Lys Ser Asp Ser Ser Asp Ser Gln Glu Gly
236      165                      170                      175
238 Asn Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg Arg Glu Arg Thr Ala
239      180                      185                      190
241 Phe Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala Glu Phe Ala His His
242      195                      200                      205
244 Asn Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile Ala Val Asn Leu Asp
245      210                      215                      220
247 Leu Thr Glu Arg Gln Val Lys Val Trp Phe Gln Asn Arg Arg Met Lys
248 225                      230                      235                      240
250 Trp Lys Arg Val Lys Gly Gly Gln Gln Gly Ala Ala Ala Arg Glu Lys
251      245                      250                      255
253 Glu Leu Val Asn Val Lys Lys Gly Thr Leu Leu Pro Ser Glu Leu Ser
254      260                      265                      270
256 Gly Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly Asp Ser Leu Ala Asn
257      275                      280                      285
259 Asp Asp Ser Arg Asp Ser Asp His Ser Ser Glu His Ala His Leu
260      290                      295                      300
263 (2) INFORMATION FOR SEQ ID NO: 3:
265     (i) SEQUENCE CHARACTERISTICS:
266         (A) LENGTH: 941 base pairs
267         (B) TYPE: nucleic acid
268         (C) STRANDEDNESS: both
269         (D) TOPOLOGY: linear
271     (ii) MOLECULE TYPE: cDNA
273     (iii) HYPOTHETICAL: NO
275     (iv) ANTI-SENSE: NO
278     (ix) FEATURE:
279         (A) NAME/KEY: CDS
280         (B) LOCATION: 33..941
283     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
285 GTCTTCTACC TGGAACCCGA AACTTGCATG CT ATG GAA CAC CCG CTC TTT GGC      53
286                               Met Glu His Pro Leu Phe Gly
287                               1                      5
289 TGC CTG CGC AGC CCT CAC GCC ACG GCG CAA GGC TTG CAC CCG TTC TCC      101

```

RAW SEQUENCE LISTING

DATE: 01/26/2006

PATENT APPLICATION: US/09/755,320

TIME: 10:10:26

Input Set : N:\Crif3\RULE60\09755320.raw.txt

Output Set: N:\CRF4\01262006\I755320.raw

290	Cys	Leu	Arg	Ser	Pro	His	Ala	Thr	Ala	Gln	Gly	Leu	His	Pro	Phe	Ser	
291			10					15					20				
293	CAA	TCC	TCT	CTC	GCC	CTC	CAT	GGA	AGA	TCT	GAC	CAT	ATG	TCT	TAC	CCC	149
294	Gln	Ser	Ser	Leu	Ala	Leu	His	Gly	Arg	Ser	Asp	His	Met	Ser	Tyr	Pro	
295		25					30					35					
297	GAG	CTC	TCT	ACT	TCT	TCC	TCA	TCT	TGC	ATA	ATC	GCG	GGA	TAC	CCC	AAC	197
298	Glu	Leu	Ser	Thr	Ser	Ser	Ser	Ser	Cys	Ile	Ile	Ala	Gly	Tyr	Pro	Asn	
299	40				45						50					55	
301	GAA	GAG	GAC	ATG	TTT	GCC	AGC	CAG	CAT	CAC	AGG	GGG	CAC	CAC	CAC	CAC	245
302	Glu	Glu	Asp	Met	Phe	Ala	Ser	Gln	His	His	Arg	Gly	His	His	His	His	
303					60					65					70		
305	CAC	CAC	CAC	CAT	CAC	CAC	CAT	CAG	CAG	CAG	CAG	CAC	CAG	GCT	CTG	CAA	293
306	His	His	His	His	His	His	His	Gln	Gln	Gln	Gln	His	Gln	Ala	Leu	Gln	
307				75				80					85				
309	ACC	AAC	TGG	CAC	CTC	CCG	CAG	ATG	TCT	TCC	CCA	CCG	AGT	GCG	GCT	CGG	341
310	Thr	Asn	Trp	His	Leu	Pro	Gln	Met	Ser	Ser	Pro	Pro	Ser	Ala	Ala	Arg	
311		90					95					100					
313	CAT	AGC	CTC	TGC	CTC	CAG	CCC	GAC	TCT	GGA	GGG	CCC	CCA	GAG	TTG	GGG	389
314	His	Ser	Leu	Cys	Leu	Gln	Pro	Asp	Ser	Gly	Gly	Pro	Pro	Glu	Leu	Gly	
315		105				110						115					
317	AGC	AGC	CCG	CCC	GTC	CTG	TGC	TCC	AAC	TCT	TCC	AGC	TTG	GGC	TCC	AGC	437
318	Ser	Ser	Pro	Pro	Val	Leu	Cys	Ser	Asn	Ser	Ser	Ser	Leu	Gly	Ser	Ser	
319	120				125						130				135		
321	ACC	CCG	ACT	GGG	GCC	GCG	TGC	GCG	CCG	GGG	GAC	TAC	GGC	CGC	CAG	GCA	485
322	Thr	Pro	Thr	Gly	Ala	Ala	Cys	Ala	Pro	Gly	Asp	Tyr	Gly	Arg	Gln	Ala	
323				140						145					150		
325	CTG	TCA	CCT	GCG	GAG	GCG	GAG	AAG	CGA	AGC	GGC	GGC	AAG	AGG	AAA	AGC	533
326	Leu	Ser	Pro	Ala	Glu	Ala	Glu	Lys	Arg	Ser	Gly	Gly	Lys	Arg	Lys	Ser	
327				155					160				165				
329	GAC	AGC	TCA	GAC	TCC	CAG	GAA	GGA	AAT	TAC	AAG	TCA	GAA	GTC	AAC	AGC	581
330	Asp	Ser	Ser	Asp	Ser	Gln	Glu	Gly	Asn	Tyr	Lys	Ser	Glu	Val	Asn	Ser	
331			170				175					180					
333	AAA	CCC	AGG	AAA	GAA	AGG	ACA	GCA	TTT	ACC	AAA	GAG	CAA	ATC	AGA	GAA	629
334	Lys	Pro	Arg	Lys	Glu	Arg	Thr	Ala	Phe	Thr	Lys	Glu	Gln	Ile	Arg	Glu	
335		185				190						195					
337	CTT	GAA	GCA	GAA	TTT	GCC	CAT	CAT	AAT	TAT	CTC	ACC	AGA	CTG	AGG	CGA	677
338	Leu	Glu	Ala	Glu	Phe	Ala	His	His	Asn	Tyr	Leu	Thr	Arg	Leu	Arg	Arg	
339	200				205						210				215		
341	TAC	GAG	ATA	GCA	GTG	AAT	CTG	GAT	CTC	ACT	GAA	AGA	CAG	GTA	AAA	GTC	725
342	Tyr	Glu	Ile	Ala	Val	Asn	Leu	Asp	Leu	Thr	Glu	Arg	Gln	Val	Lys	Val	
343				220						225					230		
345	TGG	TTC	CAA	AAC	AGG	CGG	ATG	AAG	TGG	AAG	AGG	GTA	AAG	GGT	GGA	CAG	773
346	Trp	Phe	Gln	Asn	Arg	Arg	Met	Lys	Trp	Lys	Arg	Val	Lys	Gly	Gly	Gln	
347			235						240				245				
349	CAA	GGA	GCT	GCG	GCT	CGG	GAA	AAG	GAA	CTG	GTG	AAT	GTG	AAA	AAG	GGA	821
350	Gln	Gly	Ala	Ala	Ala	Arg	Glu	Lys	Glu	Leu	Val	Asn	Val	Lys	Lys	Gly	
351			250				255					260					
353	ACA	CTT	CTC	CCA	TCA	GAG	CTG	TCG	GGA	ATT	GGT	GCA	GCC	ACC	CTC	CAG	869
354	Thr	Leu	Leu	Pro	Ser	Glu	Leu	Ser	Gly	Ile	Gly	Ala	Ala	Thr	Leu	Gln	

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/755,320

DATE: 01/26/2006

TIME: 10:10:27

Input Set : N:\Crf3\RULE60\09755320.raw.txt

Output Set: N:\CRF4\01262006\I755320.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:34 M:238 W: Alpha Fields not Ordered, Reordered [(C) CLASSIFICATION:] of (1)(vi)